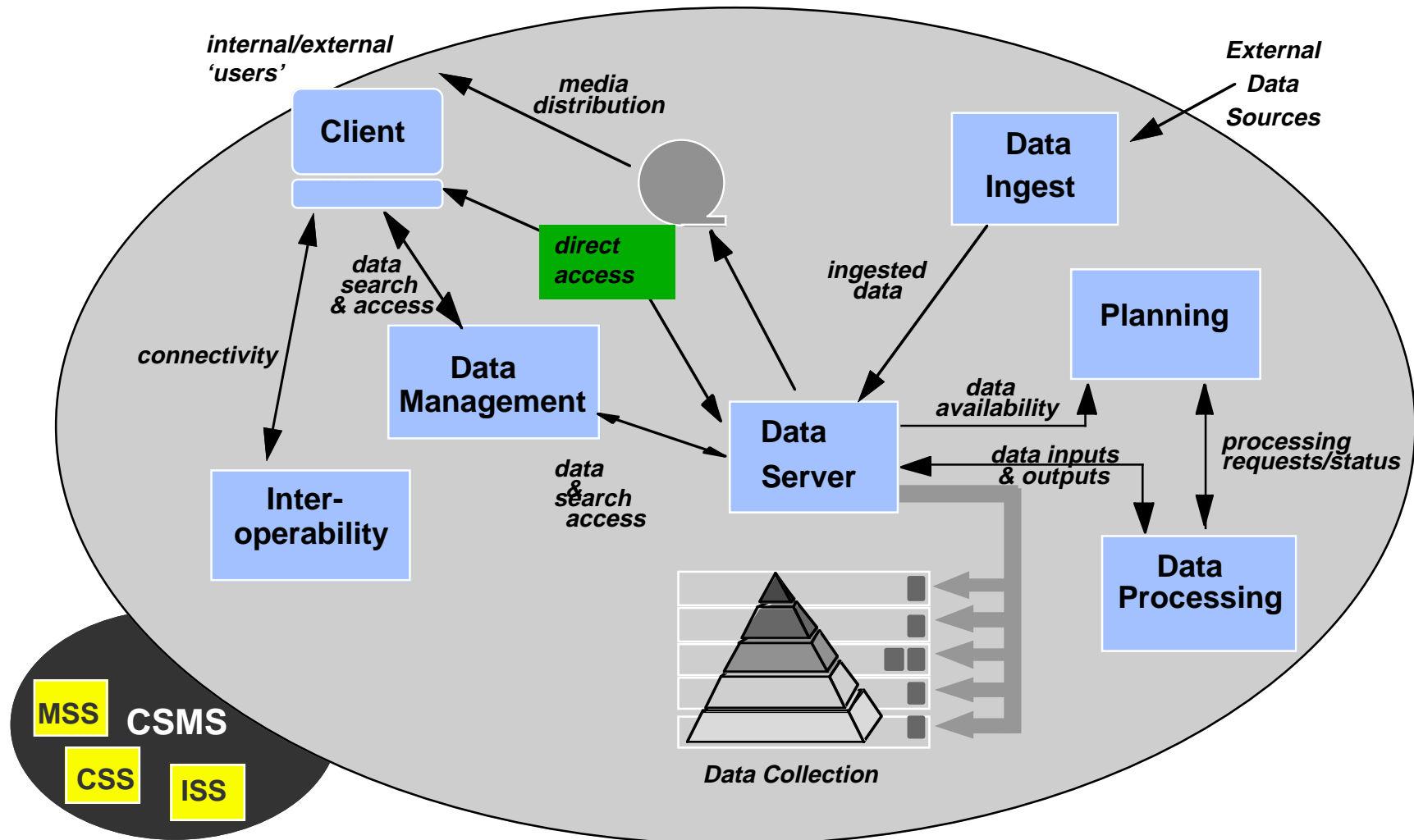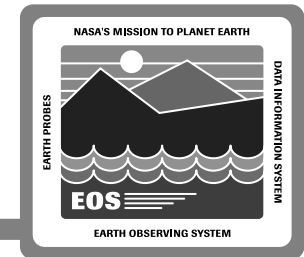# Planning and Data Processing

Karin Loya

kloya@eos.hitc.com

ECS Release A SDPS/CSMS Critical Design Review
16 August 1995

SDPS Subsystems

internal/external 'users'

Client

media distribution

direct access

External Data Sources

Data Ingest

data search & access

ingested data

Data Management

Planning

connectivity

data availability

data & search access

Data Server

processing requests/status

data inputs & outputs

Inter-operability

Data Processing

Data Collection

MSS **CSMS**

CSS   ISS

NASA'S MISSION TO PLANET EARTH
DATA INFORMATION SYSTEM
EARTH PROBES
EOS
EARTH OBSERVING SYSTEM

PDPS Presentation Road Map

Overview

- **Context**
- **Definitions**
- **PDPS High-Level Description**
    - **Functions**
    - **Design Drivers**
- **Key Design Features**
- **Major Activities Since PDR**
- **COTS Selection and Integration into Design**
- **Design Changes Since PDR**
- **Software Architecture**

PDPS Presentation Road Map (cont.)

Planning Subsystem

- **Overview**
- **Context**
- **Subsystem Decomposition**
- **Software Components**
- **Scenarios**
- **Hardware Components**
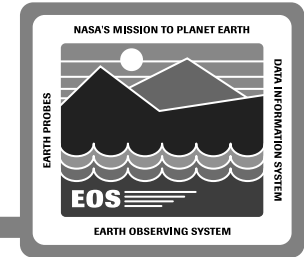- **Scalability, Evolvability**
- **Issues and Recommendations**

PDPS Presentation Road Map (cont.)

Data Processing Subsystem

- **Overview**
- **Context**
- **Subsystem Decomposition**
- **Software Components**
- **Scenarios**
- **Hardware Components**
- **Scalability and Evolvability**
- **Issues and Recommendations**

Definitions

Production Request (PR)

- **The mechanism for a Production Planner to request product(s) be generated. Will explode into multiple Data Processing Requests. A PR is contained in the PDPS Database.**

Data Processing Request (DPR)

- **Generated by Planning Subsystem from a PR. One DPR corresponds to a single Product Generation Executive (PGE) to be executed.**

Product Generation Executive (PGE)

- **Smallest entity that may be scheduled and activated by Planning and Processing. One executable or more (in combination with a script) which is designed to generate one or more standard data product(s). A PGE may be for pre-processing or post-processing.**

Definitions (cont.)
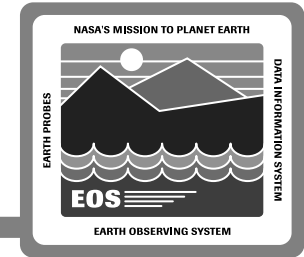
Plan

- **A timeline which embodies the processing objectives for a particular time period**
  - **a Candidate Plan is the result of "what if" planning by the Production Planner**
  - **an Active Plan is a Candidate Plan that has been selected and placed into execution by the Production Planner**

Schedule

- **Ordered set of data processing requests, as exported from Planning and imported into AutoSys; also a window into the active plan**

Ground Event

- **Non-production processing request for resources (e.g., scheduled down time)**

Definitions (cont.)

Standard Processing

- Pre-defined science software production processing which is periodic and keyed to data arrival
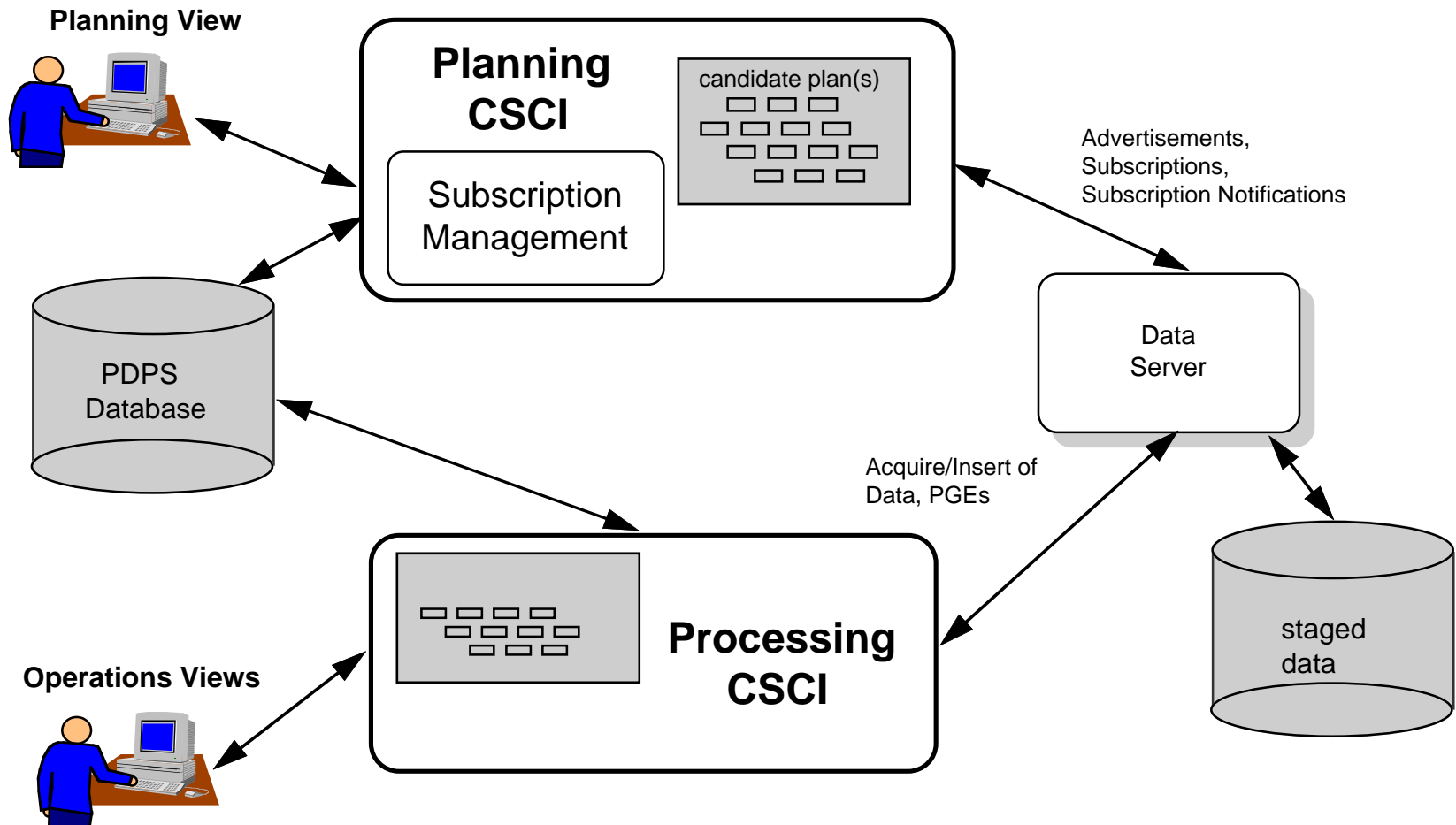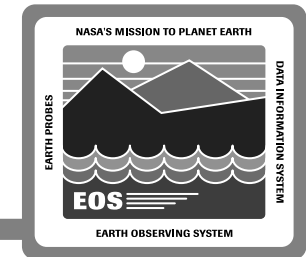
Reprocessing

- Repeat of production processing to generate replacement product(s) (limited in Release A)

On-Demand Processing

- Processing which is initiated by a user request (Release B)

Planning and Data Processing Inter-Subsystem Interfaces

Planning View

Planning CSCI

candidate plan(s)

Subscription Management

Advertisements, Subscriptions, Subscription Notifications

Data Server

PDPS Database

Acquire/Insert of Data, PGEs

Operations Views

Processing CSCI

staged data

Major Functions of Planning CSCI

- **Generates and Activates Production Plans**

- **Monitors Data Arrivals and Releases Data Processing Requests**

- **Provides Capabilities to Maintain the Plan**

- **Provides Impact Analysis Capabilities for Multiple Candidate Plans**

- **Supports Variable Planning Horizon**

Major Functions of
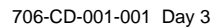# Data Processing CSCI

Performs Data Production

- **Manages data staging/destaging to/from science data processing hardware**

- **Manages resources for executing jobs**

- **Manages job execution**

Provides Capabilities for Maintaining and Controlling the Processing

- **GUIs include hierarchical job dependency diagrams, production timeline**

- **Operator alerts, alarms**

Allows "what if" Scenarios to be Examined for Current Schedule

Design Approach

NASA'S MISSION TO PLANET EARTH
EARTH PROBES
DATA INFORMATION SYSTEM
EOS
EARTH OBSERVING SYSTEM

- AHWGP Input
- Design Telecon
- workshops
- IT feedback
- Modeling input

Object Model Design

- Formal OMT
- Object Model
- Dynamic Model
- Static Model

User Feedback

- Re-use of Class Libs./Delphi

**SDS Architecture
L3, 4 Requirements
Preliminary Design
Specification**

Prototype

**Planning
and
Data Processing**

Trades

**COTS Evaluation
and Selection**

•AutoSys/AutoXpert
Selected as Scheduling
Engine

- Scheduling Engine

Design Drivers for PDPS

Data-Driven Product Generation to Increase Automation
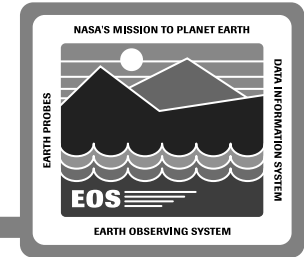
Schedule-Driven Adjustments for Flexibility

Support for activation rules identified by AHWGP

- Temporal coverage overlap
- New instance of product requires prior instance of same product
- M instances of process A run only when N instances of B complete
- Tessellation (mosaicing)   Release B
- Alternative inputs   Release B
- Conditional activation based on metadata Release B

Extensive Automation

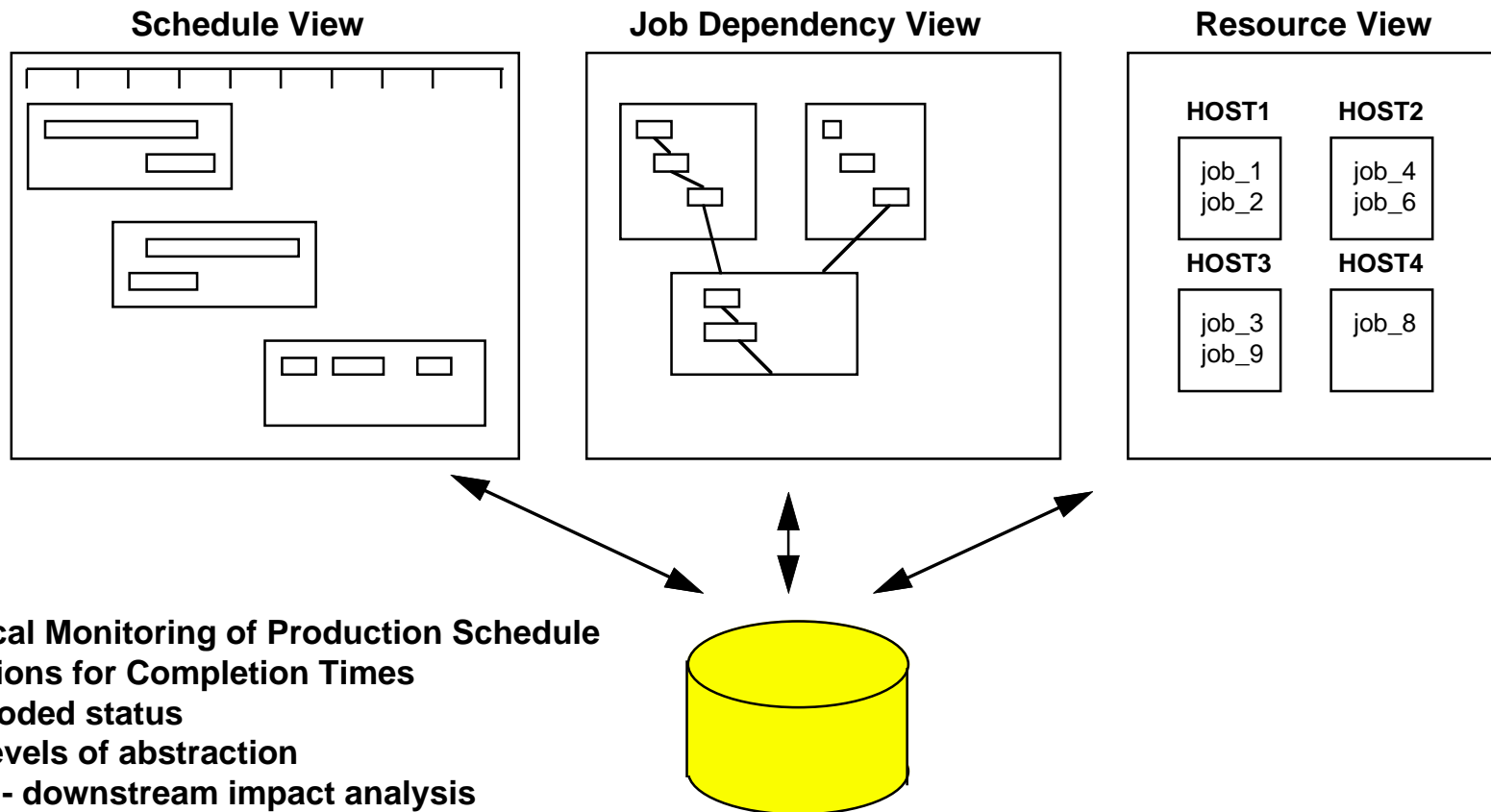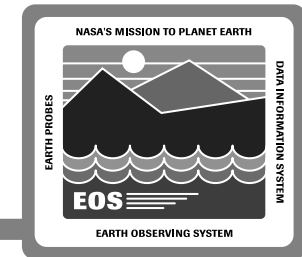**Efficient Science Data Processing Resource Utilization**

Major Activities Since PDR

Key Scheduling COTS Evaluation and Selection: Platinum AutoSys/ AutoXpert

- **Provides Graphical Representations of Schedule, Dependencies, Resources**

- **Supports Many Levels of Abstraction**

- **Offers a Sophisticated Mechanism for Tailoring Information Displayed to User Categories (ITs, Operations, Operations Management**

AutoXpert

Schedule View

Job Dependency View

Resource View

HOST1   HOST2

job_1   job_4
job_2   job_6

HOST3   HOST4

job_3   job_8
job_9

- **Graphical Monitoring of Production Schedule**
- **Predictions for Completion Times**
- **Color coded status**
- **Many levels of abstraction**
- **What If - downstream impact analysis**

AutoSys/AutoXpert Integration into PDPS

Detailed Design of PDPS

- **Integrates AutoSys/AutoXpert into Data Processing Subsystem design as scheduling engine for production processing**
  - **job management**
- **Provides elegant solution for schedule impact analysis**
- **Simplifies interfaces between planning and data processing**
- **Does not integrate AutoSys/AutoXpert into Planning Subsystem design for candidate plan creation/analysis**

AutoSys/AutoXpert Integration into PDPS (Cont)

PDPS AutoSys/AutoXpert Interface Prototyping Activity

- **To streamline AutoSys/AutoXpert integration into PDPS**
- **To identify potential problems**
- **To explore possible advantages of applicability within Planning Subsystem**
  - **COTS-based alternative for current design of one CSC**
  - **SLOC savings**
  - **Same elegant solution for "what if" analysis as for Data Processing (simulation engine)**
  - **Interoperability with scheduling**
  - **Better scalability for Release B**
  - **Better choice for evolvability**

AutoSys/AutoXpert Integration into PDPS

Anticipate Resolution of Applicability to Planning Subsystem Component by mid-September

- **Impacts across ECS assessed**
- **Impact to Release B assessed**
- **Alternate Planning Subsystem component design**
- **Delta Design Review of Planning (Workbench) CSC on September 22, 1995**

Recommended RID drafted to track design closure

Phased Implementation Plan Unaffected, as Component to be Developed in Second Phase (Begins December 1, 1995)

Other Design Changes Since PDR

Interfaces Simplified

- **MSS interface simplified (resource availability data stored within Planning Subsystem)**
- **Planning/Data Processing share common database**
- **Planning based on internal predictions instead of Data Availability Schedule**

Sybase RDBMS selected for ECS

Pre-processing CSCI Functionality Allocated to Processing CSCI and to Ingest Subsystem

PDPS Software Architecture

PDPS Application Layer

Processing Applications

Planning Applications

COTS

AI&T Applications

PDPS Persistent Data Layer
Planning Data
Processing Data
COTS schemas
 AI&T Data

SYBASE RDBMS

PDPS Database